

Article

The Molecular Basis for *Escherichia coli* O157:H7 Phage FAHEc1 Endolysin Function and Protein Engineering to Increase Thermal Stability

Michael J. Love^{1,2}, David Coombes¹, Sarah H. Manners¹, Gayan S. Abeysekera¹, Craig Billington^{2,*} 
and Renwick C. J. Dobson^{1,3,*} 

¹ Biomolecular Interaction Centre and School of Biological Sciences, University of Canterbury, Christchurch 8041, New Zealand; michael.love@pg.canterbury.ac.nz (M.J.L.); david.coombes@pg.canterbury.ac.nz (D.C.); sma345@uclive.ac.nz (S.H.M.); aagsampath@yahoo.com (G.S.A.)
² Institute of Environmental Science and Research, Christchurch 8041, New Zealand
³ Department of Biochemistry and Molecular Biology, University of Melbourne, Melbourne 3052, Australia
* Correspondence: craig.billington@esr.cri.nz (C.B.); renwick.dobson@canterbury.ac.nz (R.C.J.D.); Tel.: +64-3-351-0128 (C.B.); +64-3-364-2987 (R.C.J.D.)

Abstract: Bacteriophage-encoded endolysins have been identified as antibacterial candidates. However, the development of endolysins as mainstream antibacterial agents first requires a comprehensive biochemical understanding. This study defines the atomic structure and enzymatic function of *Escherichia coli* O157:H7 phage FAHEc1 endolysin, LysF1. Bioinformatic analysis suggests this endolysin belongs to the T4 Lysozyme (T4L)-like family of proteins and contains a highly conserved catalytic triad. We then solved the structure of LysF1 with x-ray crystallography to 1.71 Å. LysF1 was confirmed to exist as a monomer in solution by sedimentation velocity experiments. The protein architecture of LysF1 is conserved between T4L and related endolysins. Comparative analysis with related endolysins shows that the spatial orientation of the catalytic triad is conserved, suggesting the catalytic mechanism of peptidoglycan degradation is the same as that of T4L. Differences in the sequence illustrate the role coevolution may have in the evolution of this fold. We also demonstrate that by mutating a single residue within the hydrophobic core, the thermal stability of LysF1 can be increased by 9.4 °C without compromising enzymatic activity. Overall, the characterization of LysF1 provides further insight into the T4L-like class of endolysins. Our study will help advance the development of related endolysins as antibacterial agents, as rational engineering will rely on understanding mutable positions within this protein fold.

Keywords: endolysin; bacteriophage; peptidoglycan hydrolase; T4 lysozyme



Citation: Love, M.J.; Coombes, D.; Manners, S.H.; Abeysekera, G.S.; Billington, C.; Dobson, R.C.J. The Molecular Basis for *Escherichia coli* O157:H7 Phage FAHEc1 Endolysin Function and Protein Engineering to Increase Thermal Stability. *Viruses* **2021**, *13*, 1101. <https://doi.org/10.3390/v13061101>

Academic Editor: Dann Turner

Received: 20 April 2021

Accepted: 7 June 2021

Published: 9 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bacteriophages (phages) are viral parasites of bacteria and thus have bactericidal properties. Phages have been investigated for the biocontrol of bacteria in health, agriculture and food contexts [1,2]. Their high bacterial specificity and amenability to genetic engineering are advantageous properties that allows for the development of phages with optimized bactericidal activity [3,4]. Global implementation of phage-based therapy has, however, been hindered by a poor understanding of phage biology and also regulatory hurdles [5–7]. Conversely, it has been discovered that some phage-encoded proteins alone, such as endolysins, have direct antibacterial activity [8], which provides a less complex system that will be easier to regulate.

Endolysins are enzymes, encoded by virulent phages, produced at the end of the phage lytic cycle. Endolysins specifically cleave the various bonds within peptidoglycan. Degradation of the cell wall causes bacterial lysis to occur, releasing progeny phages [9]. Exogenous application of recombinant endolysins can also cause cell lysis to occur, killing

the bacteria [10]. This phenomenon has led to a growing number of investigations focused on the implementation of endolysins as antibacterial agents [11].

The application of endolysins against Gram-negative bacteria is generally hindered by the protective outer membrane that prevents access to the peptidoglycan [12]. Targeting Gram-negative bacteria using endolysins therefore relies on innovative approaches to penetrate through the outer membrane. Consequently, this technical hurdle has caused most to focus on characterizing and developing Gram-positive targeting endolysins [11,13,14], which has led to a number of endolysin-based clinical trials [15–18]. However, in recent years there have been successes using protein engineering to modify endolysins so they can penetrate through the outer membrane and subsequently cause cell lysis of Gram-negative bacteria [19,20]. Therefore, there is a need for characterization of the biophysical, cell lysis-causing and structural properties of Gram-negative targeting endolysins. These characterization studies may uncover endolysins with unique properties or may identify the best candidates for modification with outer membrane-penetrating functionality. Further, increasing the fundamental understanding of endolysins will allow for rational engineering to create novel endolysins with optimized properties for the effective killing of bacteria.

The objective of this work is to define the structural and functional properties of an uncharacterized endolysin, LysF1. The sequence of this endolysin was derived from *Escherichia coli* O157:H7 phage, FAHEc1, which was identified as a potential biocontrol agent against foodborne pathogenic strains and displays the useful attribute of high specificity for the Shigatoxigenic *E. coli* serogroup O157 [21,22]. Our characterization and comparative analysis of LysF1 provide a better understanding of the biology of FAHEc1, which may aid in its implementation as a phage-based biocontrol, and the elucidation of the sequence–structure relationship will support the modification of this class of endolysins into potent antibacterial agents.

2. Materials and Methods

2.1. Bioinformatic Analysis

The basic local alignment search tool, BLASTp (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>, accessed 10 January 2021) was used to identify sequences of Protein Data Bank (PDB) structures with similarity to LysF1 [23]. These sequences were then used in a multisequence alignment using MAFFT (<https://mafft.cbrc.jp/alignment/server/> accessed on 28 March 2021) that was rendered with Esript (<http://esript.ibcp.fr> accessed on 28 March 2021) [24,25]. The ProtParam tool from the ExPASy server was used to compute the biochemical parameters of the protein from the amino acid sequence (<https://web.expasy.org/protparam/>, accessed 10 January 2021) [26].

2.2. Expression and Purification of LysF1

The gene for LysF1 and mutants used in this study, cloned into the expression vector pET28(a)-TEV between NdeI and XhoI restriction sites, were purchased from Genscript Biotech Corporation (Piscataway, NJ, USA). This plasmid encoded the gene of interest with an N-terminal polyhistidine expression tag. BL21(DE3) cells were transformed with the plasmid and cultured in LB medium, supplemented with kanamycin ($30 \mu\text{g mL}^{-1}$), at 37°C with shaking at 180 rpm until cells reached the mid-exponential phase ($\text{OD}_{600} \sim 0.6$). Protein expression was induced with isopropyl β -D-1-thiogalactopyranoside (final concentration of 1 mM) at 37°C for 3.5 h. Cells were pelleted at $8000 \times g$ for 15 min at 4°C in a Sorvall™ Lynx 6000 Centrifuge (Thermo Fisher Scientific, Waltham, MA, USA). The pellet was resuspended in lysis buffer (20 mM sodium phosphate, 150 mM sodium chloride, 20 mM imidazole, pH 6.5) and cells were lysed for 15 min by sonication at 70%, with a 0.5 s on/off cycle using an ultrasonic processor UP200S (Hielscher, Teltow, Germany) on ice. Cell-free supernatant, containing soluble protein, was collected by centrifugation at $18,000 \times g$ at 4°C for 15 min.

LysF1 was purified by immobilized metal affinity chromatography followed by size exclusion chromatography. The cell-free supernatant was loaded onto a 5 mL HisTrap FF

column (GE Healthcare, Chicago, IL, USA) equilibrated with lysis buffer, and the column was washed with lysis buffer until a steady baseline UV reading was reached. LysF1 was then eluted in the elution buffer (20 mM sodium phosphate, 150 mM sodium chloride, 500 mM imidazole, pH 6.5). Fractions containing LysF1 were identified with SDS-PAGE analysis and then pooled. Pooled protein was spin-concentrated to 2 mL with a Vivaspin 6, 10,000 MWCO PES spin concentrator (Sartorius, Gottingen, Germany). The aggregate was pelleted at $18,000\times g$ at 4 °C for 15 min and the supernatant was loaded onto a Superdex 200 Increase 15/300 GL SEC column (GE Healthcare, Chicago, IL, USA) equilibrated with a size-exclusion buffer (20 mM sodium phosphate, 150 mM sodium chloride, pH 6.5). The purification of LysF1 to near homogeneity was visualized with SDS-PAGE analysis (Figure S1).

2.3. Turbidity Reduction Assay

The peptidoglycan degrading activity of LysF1 and the Glu88 mutants was evaluated against outer membrane-permeabilized Gram-negative bacterial cells and untreated Gram-positive *Staphylococcus aureus* ATCC 25923 cells by measuring reductions in optical density at 600 nm (OD_{600}). Gram-negative strains *E. coli* TOP10, *Salmonella typhimurium* LT2 and *Klebsiella pneumoniae* ATCC 13883 were used. Permeabilization was performed using the chloroform-Tris-HCl technique [27]. Briefly, cells were grown at 37 °C to an OD_{600} of 0.6. Cells were pelleted and resuspended in 50 mM Tris-HCl (pH 7.7) saturated with chloroform. Cells were incubated for 45 min, then washed and resuspended in 10 mM sodium phosphate buffer to an OD_{600} of 1.0. Protein (30 μ L) was combined with chloroform-treated cells (270 μ L), and the change in OD_{600} was measured over time.

The effect of temperature on LysF1 activity was assessed by incubating protein (1 μ g mL^{-1}) at each temperature for 30 min. Protein samples were then cooled at room temperature and the residual activity measured against outer membrane-permeabilized *E. coli* TOP10 cells.

The pH dependence of LysF1 activity was assessed by incubating protein with 270 μ L of *E. coli* TOP10 cells resuspended in a universal buffer (50 mM potassium chloride, 10 mM monopotassium phosphate, 10 mM trisodium citrate and 10 mM boric acid) adjusted to different pH values with NaOH and HCl.

2.4. Differential Scanning Fluorimetry

The thermal stability and denaturation of LysF1 were determined by differential scanning fluorimetry using a Quantstudio 3 Real-Time PCR system (Thermo Fisher Scientific, Waltham, MA, USA). Samples were prepared with 5 \times SYPRO[®] Orange and LysF1 and mutants at 0.8 mg mL^{-1} . Samples were measured and prepared in triplicate (25 μ L reaction volume) in a 96-well thick wall plate. Fluorescence was measured for 5 s in 0.2 °C increments from 4 °C to 95 °C.

2.5. X-ray Crystallography

Crystallization was performed by the sitting-drop vapor-diffusion method at 20 °C. Hexagonal prism-shaped crystals were produced at 20 °C in a drop which contained LysF1 concentrated to 15 mg mL^{-1} in size-exclusion buffer and mixed in a 1:1 ratio with 0.1 M sodium chloride, 1.6 M ammonium sulfate and 0.1 M sodium HEPES, pH 7.5. (G12, SG1[™] Screen HT-96 (Molecular Dimensions, Suffolk, UK)) (Figure S2). The crystal was soaked in 15% cryoprotectant (50% ethylene glycol, 50% glycerol) and mother liquor before being snap-frozen in liquid nitrogen. Diffraction data were collected on the Australian Synchrotron MX2 beamline (Melbourne, Australia). The best data set was collected to a maximum resolution of 1.71 Å. The measured data set was scaled and quality assessed using AIMLESS [28]. Data collection and refinement stats are reported in Table 1.

Table 1. X-ray data collection and refinement statistics of LysF1. Statistical values of the highest resolution shell are given in parentheses.

Data Collection Statistics	
Wavelength (Å)	0.953719
Temperature (K)	100
Detector	EIGER X 16M (Dectris)
Space group	$P 3_2 2 1$
Unit cell parameters (Å)	$a = 67.87, b = 67.87, c = 134.44$
Resolution range (Å)	44.81–1.71 (1.74–1.71)
Observed reflections	329,379 (15,760)
Unique reflections	39,743 (2042)
Mean I/σ(I)	13.1 (1.2)
CC _{1/2}	0.999 (0.531)
Completeness (%)	99.8 (96.1)
R_{merge}	0.074 (1.645)
$R_{\text{r.i.m.}}$	0.079 (1.761)
$R_{\text{p.i.m.}}$	0.028 (0.619)
Wilson B value (Å ²)	29.9
V_M (Å ³ Da ⁻¹)	2.30
Molecules per asymmetric unit	2
Solvent content (%)	46
Structure and Refinement Statistics	
R_{factor} (%)	18.4
R_{free} (%)	20.7
Number of atoms	
Macromolecules	2483
Water	299
Ligands	20
Root-mean-square deviation (r.m.s.d.)	
Bonds (Å)	0.01
Angles (°)	1.8
Average B factors (Å ²)	
Macromolecules	34.8
Water	43.6
Ligands	59.7
Ramachandran plot, residues (%)	
Favored region	99
Allowed region	1
Disallowed region	0
PDB ID	7M5I

The CCP4 suite of programs was used for molecular replacement and initial refinement and rebuilding [29]. The structure of LysF1 was solved with molecular replacement. First, a pruned and mutated template model of muramidase from *Acinetobacter baumannii* AB 5075UW prophage (PDB ID 6ET6) based on the sequence alignment with LysF1 was produced with CHAINSAW [30]. This model was then used to solve the structure of LysF1 using PHASER [31]. The native dataset of was initially calculated to be $P 3_1 2 1$; however, the solution only packed correctly in its enantiomorph, $P 3_2 2 1$. Refinement of the LysF1 structure was initially performed with REFMAC [32], and manual rebuilding of the structure was performed in COOT throughout refinement [33]. Finalization and model validation were conducted in PHENIX [34]. The figures presented here were produced with PyMOL.

2.6. Analytical Ultracentrifugation

Sedimentation velocity experiments were performed in a XL-I analytical ultracentrifuge (Beckman Coulter, Brea, CA, USA) with An-60 Ti rotor. Analytical ultracentrifugation was conducted with LysF1 at 0.32, 0.48 and 0.84 mg mL⁻¹ in 20 mM sodium phosphate,

150 mM sodium chloride, pH 6.5. Sample (380 μL) and reference (400 μL) were loaded into 12 mm double sector quartz cells with epon centerpieces. Data were obtained at 50,000 rpm at 20 °C using radial absorbance scans at 281 nm and a step size of 0.003 cm over 100 scans. Buffer viscosity, density and an estimate of the partial specific based on the sequence were determined using SEDNTERP [35]. SEDFIT was used to fit the data to a sedimentation coefficient $[c(s)]$ distribution model or mass distribution $[c(M)]$ [36]. The residuals were inspected to assess the quality of the data (Figure S3)

2.7. Small-Angle X-ray Scattering

Small-angle x-ray scattering data (SAXS) were collected at the Australian Synchrotron (Melbourne, Australia) on the SAXS/WAXS beamline equipped with a Pilatus 1M detector (170 \times 170 mm, effective pixel size 172 \times 172 μm). A detector distance of 1600 mm, which produced a q range of 0.006 to 0.5 \AA^{-1} . 50 μL of LysF1 at a concentration of 10.43 mg mL^{-1} , was injected onto an inline Superdex S200 increase 5/150 GL SEC column (GE Healthcare, Chicago, IL, USA), equilibrated with 20 mM sodium phosphate, 150 mM sodium chloride, pH 7.4 supplemented with 0.1% (w/v) sodium azide at a flow rate of 0.3 mL min^{-1} . Scattering data were collected at one second exposure (wavelength = 1.0332 \AA) over 400 frames using a 1.5 mm glass capillary at 25 °C. Coflow SAXS was used to minimize sample dilution and maximize the signal-to-noise ratio.

Analysis was performed using the ATSAS program suite [37]. PRIMUS was used for Guinier analysis and for producing the Kratky plot [38]. GNOM was used to calculate the pairwise distance distribution function $P(r)$ via indirect Fourier transformation [39]. For a useful comparison of the experimental and computed scattering, a full model of LysF1 needed to be used. The full N-terminal His tag and missing side chains, not modelled in the final crystal structure due to missing electron density, were generated with Modeller using the Chimera interface [40–42]. CRY SOL was used to generate the theoretical scattering curve of the full model to compare with the experimental scattering [43].

2.8. Circular Dichroism

Circular dichroism measurements were performed with a J-815 spectrophotometer (JASCO Corporation, Tokyo, Japan). Wavelength scans between 200–240 nm were recorded on purified samples of LysF1 and mutants in 20 mM sodium phosphate buffer (pH 7.0) at a protein concentration of 0.26–0.30 mg mL^{-1} in a 1 mm path-length quartz cuvette in 0.2 nm steps.

3. Results and Discussion

3.1. Protein Sequence Analysis Suggests LysF1 Is a Glycoside Hydrolase Family 24 Member

Bioinformatic analysis of LysF1 was performed to identify the protein family of the endolysin and the potential catalytic mechanism. We first analyzed the primary sequence of LysF1 with ProtParam [26], identifying that the protein is 154 amino acids, with a theoretical isoelectric point (pI) of 8.89. BLASTp analysis with non-redundant protein sequences indicated that LysF1 belongs to the lysozyme-like glycoside hydrolase family 24 (GH24). This group contains phage endolysins, such as the well-studied T4 lysozyme (T4L), and bacterial autolysins, which cleave the β -1,4 glycosidic bond between *N*-acetylmuramic acid and *N*-acetylglucosamine in peptidoglycan [23,44]. A high degree of similarity to the subfamily GH24v (viral-type) sequence signature was observed [44], indicating that these conserved residues play a structural or functional role in LysF1. Several similar endolysin sequences (above 90% amino acid similarity) were identified from *Escherichia*, *Salmonella* and *Shigella* phages. This high level of similarity indicates that LysF1 will likely share common features with these predicted proteins and may provide insight into these yet uncharacterized endolysins.

To further assess the structural similarity that is generally conserved in this group of endolysins, we ran a protein DELTA-BLAST algorithm against those proteins with experimentally determined structures stored in the Protein Data Bank [23,45]. Six unique endolysin

sequences with similarity to LysF1 were identified; *A. baumannii* phage lysozyme (43.2% identity, 95% query cover) [46], DLP12 endolysin (36.7% identity, 97% query cover) [47], P22 lysozyme mutant L86M (36.2% identity, 96% query cover) [48], P21 lysozyme (26.8% identity, 89% query cover) [49], P1 lysozyme (20.4% identity, 97% query cover) [50] and T4L (20.1% identity, 76% query cover) [51]. We performed a multisequence alignment with MAFFT to identify conserved and catalytically important residues in LysF1 (Figure 1).

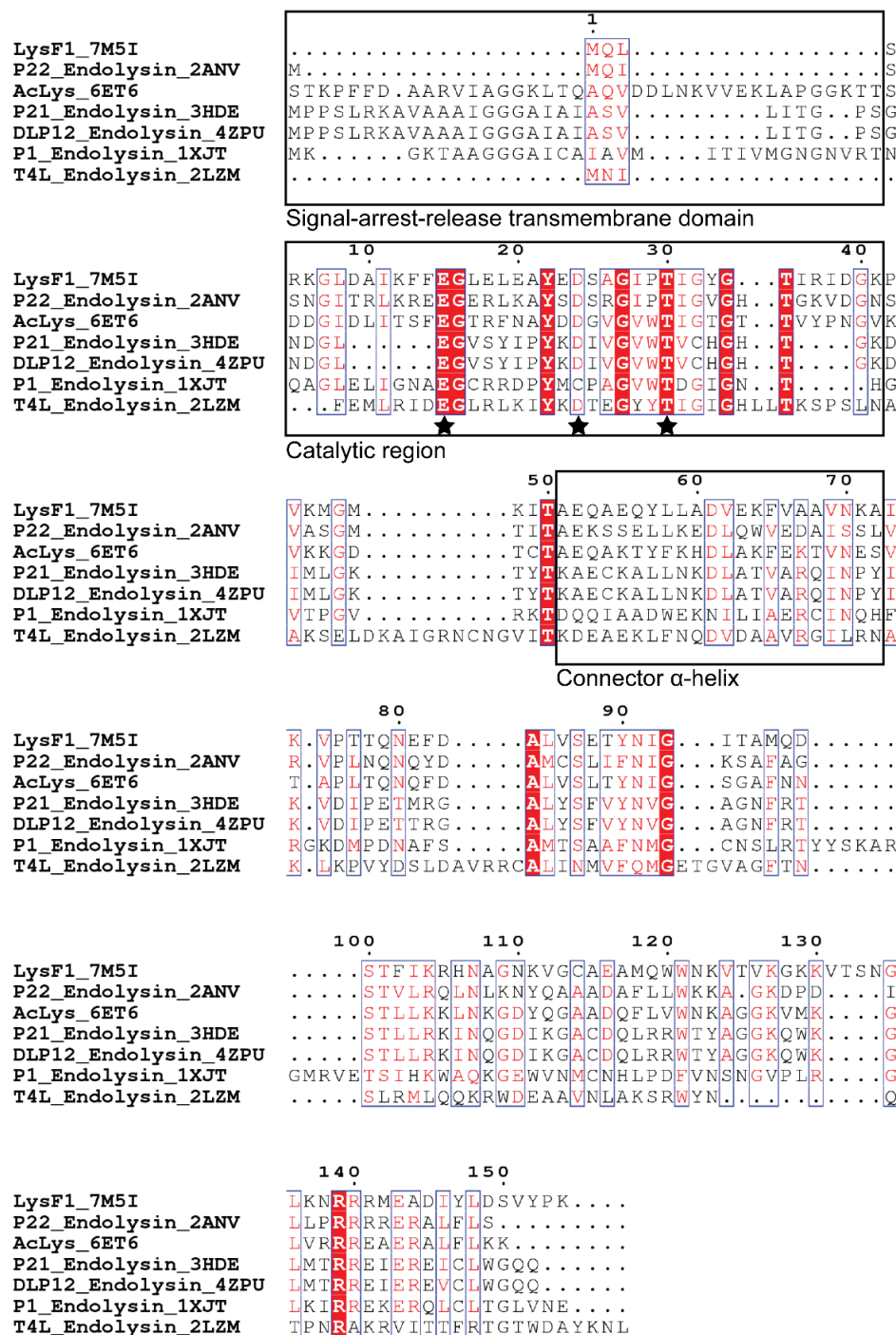


Figure 1. Sequence comparison of LysF1 with six closely related endolysins. Structural features of these endolysins are marked with black boxes and symbols. Putative catalytic residues are marked with a star symbol. Red boxes indicate a conserved residue. Red letters indicate similarity within a group and blue frames indicate similarity across a group.

Proteins that belong to the GH24v family contain a catalytic triad, consisting of a glutamate, aspartate or cysteine, and a threonine [49,52,53]. The multisequence alignment reveals that LysF1 contains these conserved residues with Glu15, Asp24 and Thr30. The conservation of these key residues suggests that LysF1 likely performs peptidoglycan degradation in a similar mechanism to T4L, the archetype of the GH24v family.

The sequence of LysF1 offers valuable information on the lytic mechanism of FAHEc1. In the canonical phage-mediated cell lysis mechanism, endolysins accumulate fully folded in the cytoplasm. Then, at the correct time, holin-mediated pore formation occurs in the plasma membrane, allowing endolysins to gain access to and hydrolyze the peptidoglycan, causing cell lysis [9,54]. An alternative lytic mechanism has recently been elucidated, where endolysins with signal peptides, signal-arrest-release regions, are transported across the plasma membrane but remain embedded in the membrane, unfolded and inactive [47,49,50]. Lysis is triggered when pin holins punch a hole in the membrane, causing the collapse of proton motive force, releasing the embedded endolysins allowing them to fold correctly and cause cell lysis [54]. Sequence comparison shows that there is no N-terminal extension on LysF1, suggesting no N-terminal signal-arrest-release region, similar to T4L and the endolysin from P22. Therefore we predict that, as LysF1 is not a signal-arrest-release endolysin, FAHEc1 lyses cells via a canonical lytic mechanism, similar to the T4 phage [55].

The sequence variation between other available structures of similar endolysins meant that we wanted to investigate our polypeptide sequence further to understand the structural implications of this diversity.

3.2. LysF1 Is an Endolysin That Causes Lysis of Outer Membrane-Permeabilized Gram-Negative Bacteria

Endolysin activity of LysF1 was tested by incubation with Gram-negative bacterial cells and monitoring optical density. *E. coli* cells were treated with chloroform to chemically remove the outer membrane layer and expose the peptidoglycan, which is the substrate of endolysins [14]. Addition of LysF1 to the treated *E. coli* cells produced a decrease in optical density consistent with the degradation of the peptidoglycan layer and cell lysis (Figure 2A). Addition of LysF1 to untreated cells did not reduce the optical density, demonstrating that LysF1 can only access the peptidoglycan substrate once the outer cell membrane is removed (Figure 2A). This behaviour has been reported with other Gram-negative endolysins and is a significant technical barrier to the therapeutic use of endolysins against Gram-negative bacteria [56–58].

To investigate the specificity of LysF1, cell lysis was measured with *S. typhimurium*, *K. pneumoniae* and *S. aureus*. Addition of LysF1 with chloroform-treated *S. typhimurium* and *K. pneumoniae* produced a decrease in optical density consistent with cell lysis (Figure 2B). *E. coli*, *S. typhimurium* and *K. pneumoniae* share the same peptidoglycan chemotype (A1 γ), but are more diverse at the protein level [59,60]. No cell lysis of *S. aureus* was observed (Figure 2B) despite the conservation of the targeted $\beta(1-4)$ glycosidic bond [59]. This indicates that the differences in the peptidoglycan composition of *S. aureus* and the A1 γ -type, such as the peptide stem or the interpeptide bridge, are important aspects of substrate recognition or degradation by LysF1. Therefore, we predict that the LysF1 endolysin is specific for the peptidoglycan substrate alone and not specific to a host species or strain. This means that LysF1 has a broader spectrum of specificity than FAHEc1, which displays near complete specificity for *E. coli* serogroup O157 [21].

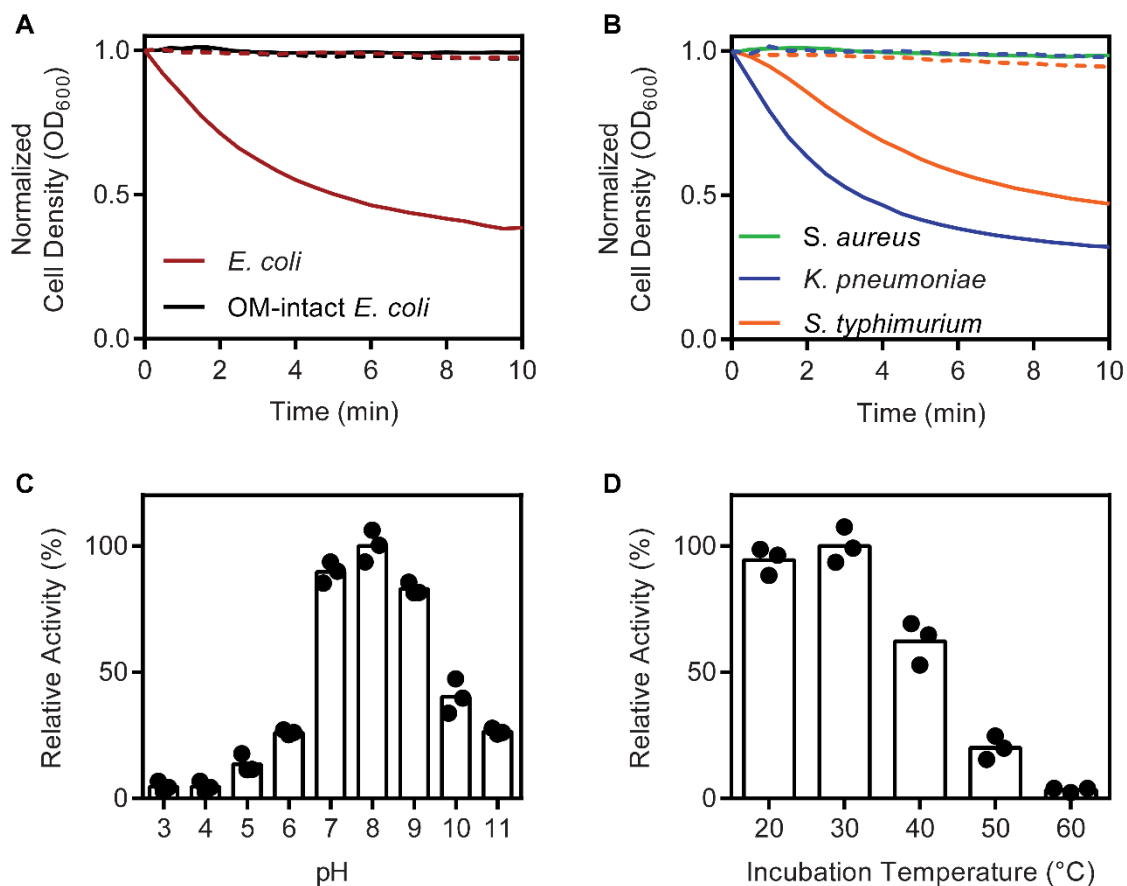


Figure 2. LysF1 is an active endolysin. (A) The incubation of LysF1 with outer membrane-permeabilized *E. coli* (red) resulted in a decrease in optical density (OD₆₀₀) indicating lysis is occurring as a result of peptidoglycan degradation. Lysis was not observed when LysF1 was incubated with *E. coli* that had intact outer membranes (OM) (black). (B) LysF1 was also active against outer membrane-permeabilized *K. pneumoniae* (blue) and *S. typhimurium* (orange). LysF1 was not active against *S. aureus* (green) with no reduction in optical density measured. Dashed lines show the optical density of each cell suspension without the addition of protein. (C) The pH dependence of LysF1 activity measured against cells resuspended in different pH buffers. (D) The residual activity of LysF1 measured after 30 min incubation at each temperature. Activity is plotted relative to the highest measured activity.

Temperature and pH uniquely affect the cell lysis-causing ability of different endolysins. Therefore, we investigated these properties of LysF1. The pH dependence of LysF1 activity was measured using cells resuspended in different buffers with a defined pH. LysF1 was able to cause cell lysis across the measured pH range 3–11 with the highest activities observed between pH 7–9 (Figure 2C). Next, we investigated the impact of temperature on LysF1 activity. Protein was incubated at set temperatures for 30 min, cooled to room temperature and then the turbidity reduction activity was measured. A decrease in activity was observed after incubation at 40 °C and only 3% of activity was retained after incubation at 60 °C (Figure 2D). LysF1 displays less temperature resistance than other endolysins, such as *Salmonella* phage endolysins Lys68 [61] and Gp110 [57], which both exhibited a non-significant decrease in activity after incubation at 60 °C.

3.3. The Structure of LysF1

To better understand the mechanism by which LysF1 functions, we crystallized LysF1 and solved the structure to a maximum resolution of 1.71 Å. We used molecular replacement to phase the structure, exploiting the *A. baumannii* phage lysozyme, AcLys, (PDB ID 6ET6) crystal model as a template. The structure was refined to a final R_{free} of 20.7% and shows

good geometry with no residues found in the disallowed region of the Ramachandran plot (Table 1), suggesting that our structure is of high quality.

The structure contains two protein molecules in the asymmetric unit each comprising the full endolysin sequence derived from FAHEc1 (residues 1–154) and two phosphates (Figure 3A). On chain A, only six residues of the N-terminal expression tag are also modelled while four are modelled on chain B. Residues are numbered based on the native structure (1–154). Electron density was well defined for most residues except for the side chains of Arg37, Asn39, Lys41 and Lys154 of chain A; these residues were therefore stubbed in the final model. Poor backbone density was found for Asn39–Gly40 of chain A; we predict that this could be caused by flexibility of the region. Chain A and B are similar with an average root mean square deviation (r.m.s.d.) of 0.334 (144 atoms used in the calculation). Further structure analysis is described using chain A.

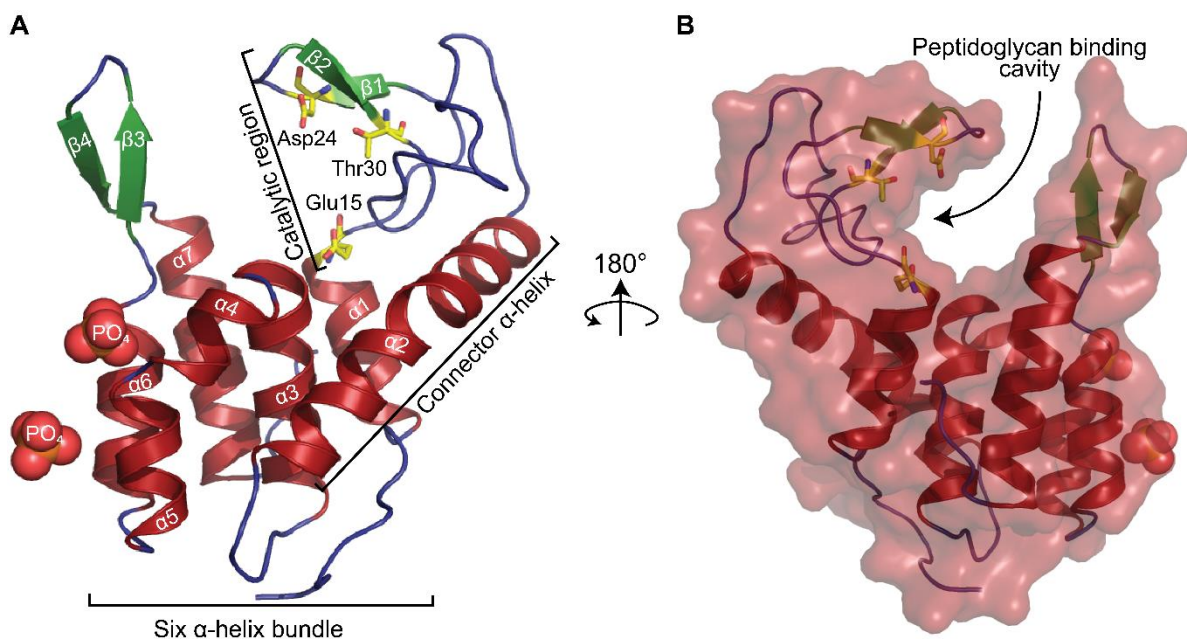


Figure 3. The monomeric structure of LysF1. Alpha helices in red, beta sheets in green and random coils in blue. Catalytic residues in yellow. (A) Cartoon representation of a monomeric structure of LysF1. (B) Surface representation overlaid upon the cartoon representation, showing a cavity in the catalytic region where peptidoglycan could bind.

The completed model reveals that the overall architecture of LysF1 consists of three regions, a catalytic loop region (Glu15–Lys41), a connector helix ($\alpha 2$, Ala51–Lys71) [47] and a C-terminal α -helix bundle ($\alpha 3$ – $\alpha 7$, Gln79–Asp149) (Figure 3A). The protein exhibits a cleft between the catalytic loop region and the α -helix bundle where peptidoglycan is proposed to bind (Figure 3B). The catalytic loop region contains a β -hairpin ($\beta 1$ – $\beta 2$), a structural element that is conserved across the glycoside hydrolase lysozyme super family and contains predicted catalytic residues Asp24 and Thr30 [44]. A putative catalytic residue Glu15 is positioned at the C-terminal end of $\alpha 1$ facing into the cleft. The catalytic loop region and C-terminal helix bundle are connected by a long α -helix, termed a ‘connector helix’, which acts as a structural backbone of the protein (Figure 3A) [47].

3.4. The Monomeric LysF1 Solution Structure Is Consistent with the Crystal Structure

The model determined from x-ray crystallography can occasionally misrepresent the true biological protein structure due to crystallization artefacts; in particular, the rigidity of a protein lattice can perturb protein dynamics. For example, the crystal structure contains two molecules in the asymmetric unit of LysF1 that interact, which could represent a weak biological interface or could have formed as an artefact of crystallography. Therefore, we

investigated the structure of LysF1 in solution using analytical ultracentrifugation and small angle x-ray scattering.

Analysis of the structure with PDBePISA (<http://pdbe.org/pisa/>, accessed on 11 March 2021) [62] shows that there is only a small interface area (79.2 Å²) between the chains and no strong interactions exist, indicating that LysF1 is unlikely to form a biologically relevant dimer structure. However, the endolysin AcLys, which has a similar protein fold and 43.2% sequence identity, was reported to exist in a monomer–dimer equilibrium in solution [46]. Therefore, we probed the quaternary structure of LysF1 in solution using analytical ultracentrifugation. Sedimentation velocity experiments were performed at three different protein concentrations (0.32, 0.48 and 0.84 mg mL⁻¹) to assess the potential equilibrium of higher order oligomeric states formation. When the data were fitted to a continuous sedimentation coefficient distribution $c(s)$ and a continuous mass distribution $c(M)$, a single peak is observed at 1.7 S and 20.2 kDa (Figure 4A) which is consistent with LysF1 existing as a monomer in solution. The frictional ratio (f/f_0) of 1.37 also suggests LysF1 is asymmetric in solution. These results confirm that the two interacting LysF1 molecules within the crystal were an artefact of protein crystallization.

Next, we validated our crystal structure compared with the solution structure using SAXS. The SAXS data statistics are summarized in Table 2. A theoretical SAXS scattering was generated from the x-ray crystal structure, with the expression tag modelled using Modeller [41], and fit to the experimental scattering curves using CRY SOL (Figure 4B) [43]. The Chi-squared value (χ^2) of 1.0702 indicates that the LysF1 structure is a good representative model of the structure in solution. Assessment of the experimental scattering in a low q region of the Guinier plot shows that the collected data were of high quality (Figure 4B, inset). The linearity of the Guinier plot signifies the homogeneity and monodispersity of the sample; nonlinearity would indicate interparticle interference or aggregation [63].

Table 2. Data analysis from the LysF1 SAXS experiment.

SAXS Data Analysis	
$I(0)$ (cm ⁻¹)	0.021
R_g (Å)	18.53
$P(r)$ analysis	
$I(0)$ (cm ⁻¹)	0.02
R_g (Å)	18.95
D_{max} (Å)	68.39
Porod volume (Å ⁻³)	27,014.90
Molar mass from Porod Volume (Da)	15,475
Calculated monomeric from sequence ^a (Da)	19,575.25

^a <https://web.expasy.org/protparam/>. Accessed 10 January 2021.

The pair-distance distribution function was determined using indirect Fourier transformation using GNOM [39]. The maximum dimension of the scattering particle (D_{max}) was calculated to be 68.39 Å (Figure 4C). The peak is largely symmetrical between 0–40.00 Å, but exhibits a shoulder peak from 40.00–68.39 Å. This suggests that the protein in solution is mostly globular, but with some elongation [64]. The Kratky plot, calculated from the scattering, is also indicative of the shape in solution (Figure 4D). The profile of the bell-shaped curve confirms that LysF1 is folded in solution; however, the plateau above zero suggests some flexibility [65]. The molecular weight estimated from the SAXS scattering profile of 15.5 kDa is consistent with the calculated monomeric weight of LysF1.

Taken together, sedimentation velocity experiment and SAXS data confirm the validity of the LysF1 structure from the x-ray crystallography experiment and further confirm that LysF1 exists as a monomer in solution.

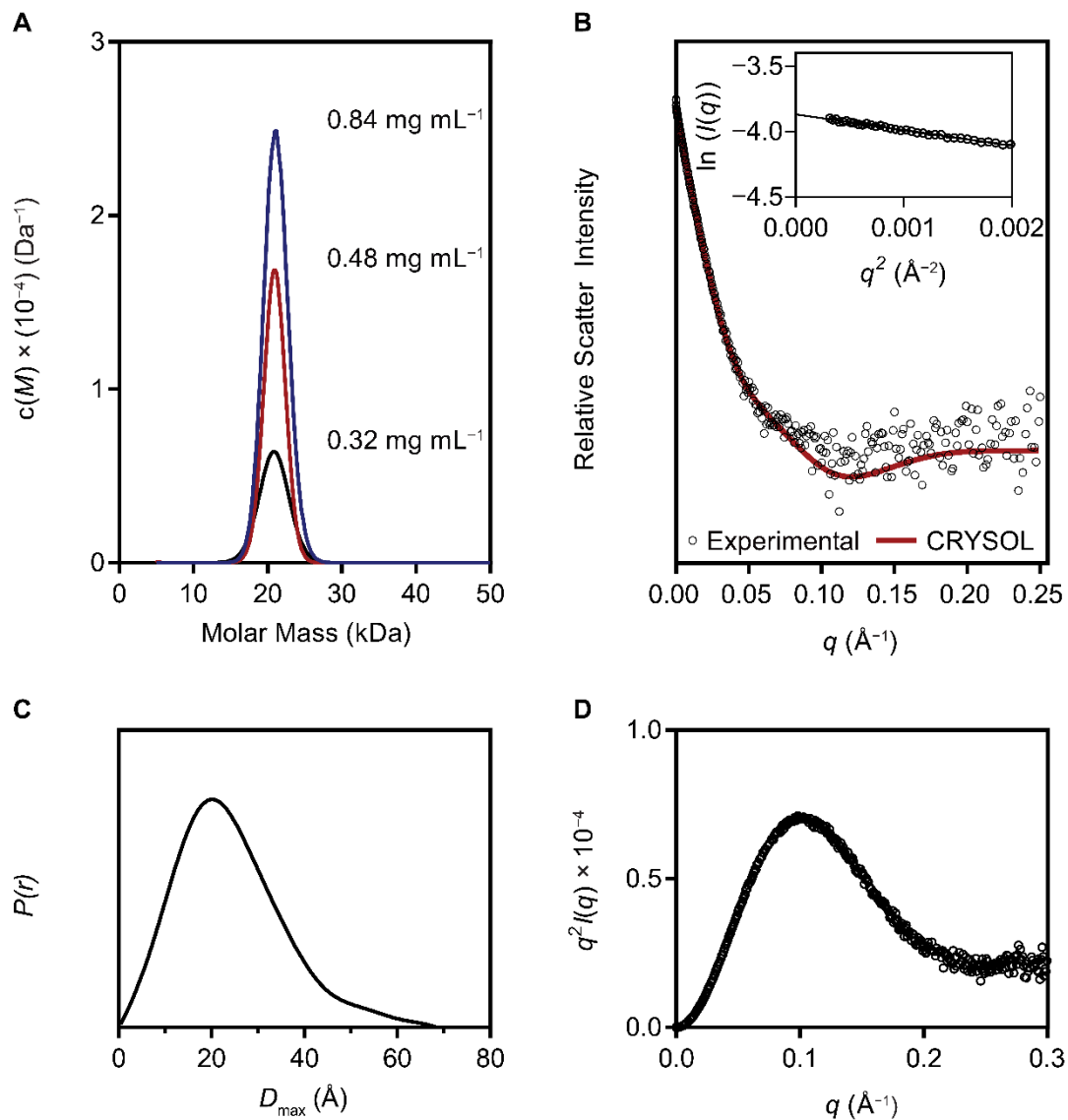


Figure 4. Solution structure of LysF1. (A) Sedimentation velocity profile of LysF1 displays a single peak at 20.2 kDa when fit to a continuous molar mass distribution model. This correlates to the molecular weight of a monomeric LysF1. (B) The biological relevance of the crystal structure is confirmed by a comparison of theoretical scattering of the LysF1 crystal structure to experimentally determined LysF1 having high similarity. The insert shows the linearity of the low- q range, indicative of high-quality data. (C) The pairwise distance distribution $P(r)$ function determined the maximum interparticle distance was 68.39 Å. (D) Kratky plot displays a mostly bell-shaped profile, which shows the protein is folded. Deviation from the baseline suggests some protein flexibility.

3.5. LysF1 Catalyzes Peptidoglycan Degradation in a T4L-like Mechanism

Our bioinformatic analysis indicated that LysF1 belongs to the group of T4L-like endolysins that utilize a well-established mechanism involving a conserved catalytic triad [47,66]. In structurally characterized T4L-like endolysins, the catalytic domain typically contains a triad of glutamate, aspartate and threonine (Table 3). However, some variation has been observed with cysteine instead of aspartate in the P1 lysozyme [50]. Interestingly, the T4L mutant Asp20Cys also retained activity [67]. The triad is arranged in a motif of Glu(X)₈(Asp/Cys)(X)₅Thr, which has been observed in all published structures from this family of endolysins. LysF1 contains this conserved catalytic triad (Figure 5A), and comparison with T4L reveals that spatial orientation is also conserved (Figure 5B).

Table 3. Structurally characterized T4L-like endolysins.

Endolysin	Putative Key Catalytic Residues			PDB ID	Reference
LysF1	Glu15	Asp24	Thr30	7M5I	This work
T4L	Glu11	Asp20	Thr26	2LZM	[51]
P1 endolysin	Glu42	Cys51	Thr57	1XJT	[50]
P22 endolysin	Glu16	Asp25	Thr31	2ANX	[48]
P21 endolysin	Glu35	Asp44	Thr50	3HDE	[49]
DLP12 endolysin	Glu35	Asp44	Thr50	4ZPU	[47]
AcLys	Glu64	Asp73	Thr79	6ET6	[46]

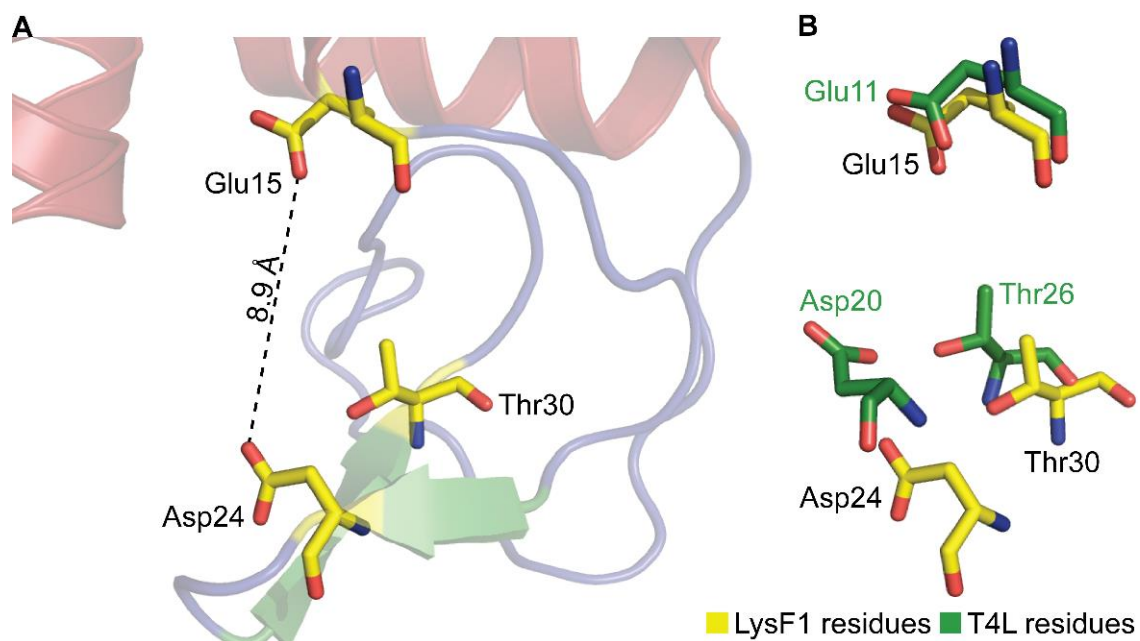


Figure 5. The catalytic site of LysF1. (A) The proposed catalytic site of LysF1 identified by sequence similarity which shows that residues Glu15 and Asp24 are positioned at an appropriate distance for the hydrolysis of the peptidoglycan. (B) Overlay of the catalytic residues of LysF1 (yellow) with T4L (green) shows the conservation of these key residues, suggesting LysF1 catalyzes the degradation of peptidoglycan using a similar mechanism as T4L.

The catalytic mechanism of peptidoglycan hydrolases targeting the β -1,4 glycosidic bond between *N*-acetylmuramic acid and *N*-acetylglucosamine can be characterized by whether the anomeric center of the cleaved peptidoglycan glycoside is retained or inverted [68]. The distance between the catalytic glutamate and aspartate is a distinguishing feature of each mechanism, with an average distance of 4.3–5.9 Å observed in enzymes that retain the anomeric center and 7.2–9.5 Å typically observed when the anomeric center is inverted [47,69,70]. T4L inverts the anomeric center during catalysis [71]. The catalytic residues of LysF1 are positioned 8.9 Å apart, consistent with T4L-like endolysins, which would accommodate the peptidoglycan and water molecule between the side chains for catalysis (Figure 5A) [47]. Therefore, the catalytic site of LysF1 is consistent with a peptidoglycan hydrolase that inverts the anomeric center of the peptidoglycan glycoside during catalysis.

3.6. LysF1 Retains T4L-Like Architecture with Context Important Mutations

We have classified LysF1 as a T4L-like endolysin based upon sequence analysis (Figure 1) and inspection of the catalytic site (Figure 5B). However, LysF1 only has 20.1% sequence identity to T4L. Therefore, we performed a comparative structural analysis to further understand the sequence–structure relationship (Figure 6).

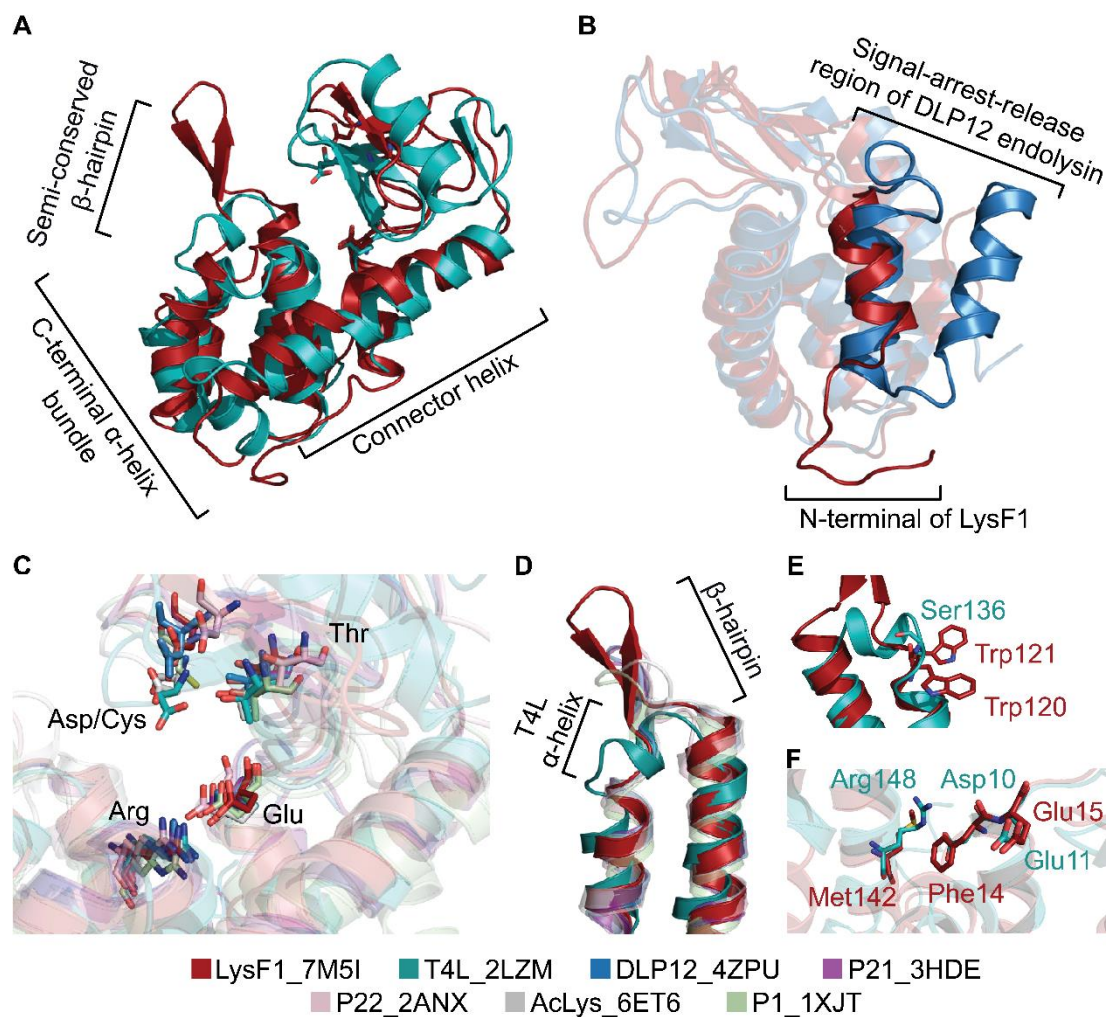


Figure 6. Comparison of LysF1 with related T4L-like endolysin structures. (A) An overlay of the cartoon model of LysF1 (red) with T4L (teal), as a representative of the T4L-like endolysin family, illustrates the conservation of the protein fold. (B) Overlay of the N-terminals of LysF1 with signal–arrest–release endolysin DLP12 (blue) shows that LysF1 does not contain the α -helix characteristic of signal–arrest–release endolysins. (C) A comparison of the active site residues highlights the conservation of Glu, Thr and Asp/Cys in terms of sequence and spatial orientation. High conservation of Arg, which is proposed to position the Glu through the formation of a salt bridge, indicates its importance. (D) The C-terminal bundle of this protein fold contains two α -helices that are connected via a β -hairpin of varying sizes in all compared structures except T4L which contains an α -helix. (E) The equivalent residue of Ser136 in T4L is a bulky tryptophan. (F) The salt bridge forming Asp10 in T4L is replaced with a hydrophobic phenylalanine residue in LysF1.

The structure of LysF1 was superimposed with T4L as a representative of the architecture observed in this protein family (Figure 6A). Despite dissimilarity in sequence identity, the superimposition of the structures revealed that the protein architecture of LysF1 is conserved (A r.m.s.d. of 2.76 was reported by RaptorX structure alignment server [72,73]). Both structures contain two domains: an α -helix bundle domain and a catalytic domain with a connector helix extending from the base of the catalytic region to beside the bundle, connecting the two domains. This architecture is observed across all structures classified as T4L-like endolysins (Figure S4).

The earlier bioinformatic analysis identified that LysF1 does not encode an N-terminal signal-arrest-release region based upon sequence analysis (Figure 1). Consistently, comparison of the N-terminal structures of LysF1 and signal–arrest–release endolysin, DLP12, confirms that LysF1 does not contain an additional α -helix that is important for insertion into the inner membrane in signal–arrest–release endolysins (Figure 6B) [47].

Comparisons of the catalytic region shows the spatial orientation of catalytic residues is conserved across the aligned endolysin structures (Figure 6C). An arginine extends from the α -helix bundle into the catalytic region forming a salt bridge to stabilize the position of the catalytic glutamate in all structures (Figure 6C) [47].

There are key residues that appear conserved across this family of proteins. Babu et al., [47] identified that in T4L, Gly12 (Gly15, LysF1 equivalent residue), Tyr18 (Tyr22), Gly23 (Gly27), Gly30 (34), Thr59 (Thr50) and Gly107 (Gly93) are conserved in endolysins with a T4L-like fold in sequence and in spatial orientation. These residues are conserved in both LysF1 and the recently characterized AcLys [46].

While there is a general conservation of protein architecture across the group, a structural difference is observed in the presence of a β -hairpin turn in the C-terminal α -helix bundle. In LysF1, a β -hairpin turn connects $\alpha 6$ and $\alpha 7$. A β -hairpin of varying sizes is seen between equivalent helices in DLP12, P21 and P1 endolysins; T4L has a small α -helix in the same position. Conversely, this region on P22 appears to be a long loop region with no α -helix or β -strand elements (Figure 6D).

The preservation of this protein architecture (Figure 6A) despite differences in the sequences (Figure 1), and the presence of numerous mutations at most of the residue positions, indicates that the fold is robust. The tolerance of T4L, the archetype of this fold, to mutation was previously shown by Rennell et al., [74]. In their study, they introduced 13 different mutations into each amino acid position and measured an approximate activity. Interestingly, point mutations that resulted in suppression of T4L activity are found in LysF1. For example, Ser136 of T4L is positioned in the small α -helical linker region between two hydrophobic core α -helices and was generally intolerant of substitution by a larger residue such as Phe and Tyr [74]. Trp120 and Trp121 were found in the similar position of LysF1 (Figure 6E). These larger residues are accommodated by other differences in the surrounding amino acid sequence that subsequently fold into different secondary structures. LysF1 contains the β -hairpin turn instead of the α -helix present in T4L. Further, a buried salt bridge between Asp10 and Arg148 in T4L was suggested to be important for stabilizing the catalytic Glu11 [74]. Asp10 is highly conserved across T4L-like sequences [52]. In LysF1, an uncharged phenylalanine residue (Phe11) is positioned into the α -helix bundle in the equivalent position but a single Glu10Phe mutation resulted in suppression of activity in T4L [74]. These observations illustrate the importance of considering the context-dependence of each mutation. A mutation may affect the phenotype of nearby residues and mutations [75]. For example, for Phe11 in LysF1 to be accommodated, there must be surrounding mutations that have coevolved when compared with the sequence of T4L. Further comparative mutagenic studies between LysF1, T4L and related endolysins would be useful to understand the role of coevolution and the impact of context-dependent mutations on the evolutionary trajectory towards this protein fold and function.

3.7. Improving the Function and Stability of LysF1 Through Mutation of the Hydrophobic Core

Inspection of the LysF1 crystal structure revealed the unusual appearance of a charged side chain (from Glu88) positioned within the hydrophobic core of the α -helix bundle (Figure 7A). Charged residues that are buried into hydrophobic regions are usually destabilizing for the protein fold but their presence is often found to have functional or catalytic importance [76,77]. Further, equivalent residues observed in other available structures were all hydrophobic. In T4L, Met102 occupies a similar spatial orientation as Glu88 (Figure 7A). Point mutations of Met102 were generally tolerated when replaced with hydrophobic residues, but residues with charged side changes resulted in loss of activity [74,78]. The T4L mutation Met102Lys was one of the most destabilizing single point mutations for the protein, with a decrease in thermal stability of 20.3 °C compared with the wild type [79,80]. Additionally, introduction of a glutamate at this position was reported to be inactivating [74]. We therefore hypothesized that LysF1 could have improved functionality through mutation of Glu88.

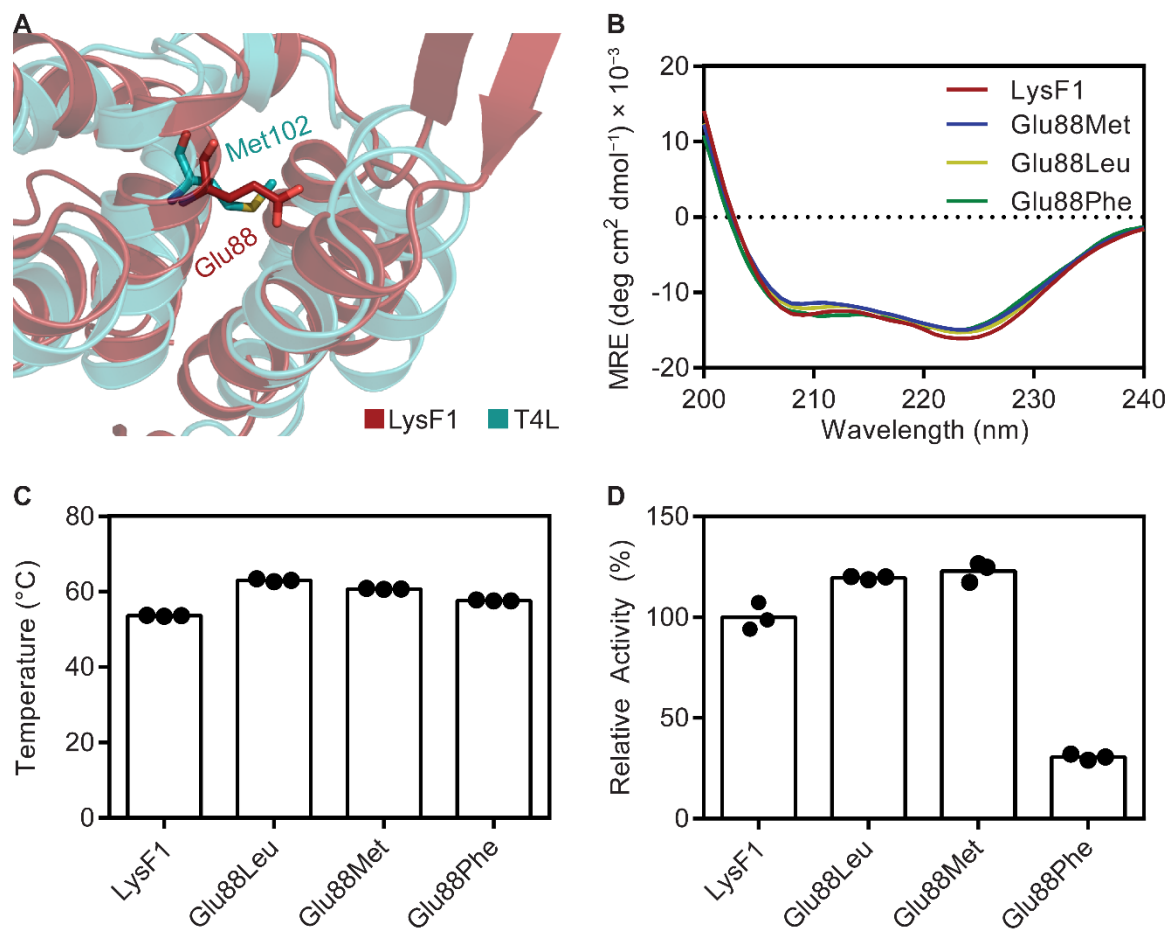


Figure 7. Thermal stability and relative activity of mutants of the charged core. (A) A magnified view of the C-terminal α -helix bundle with the crystal structures of LysF1 (red) overlaid with T4L (teal) and the residue of interest, Glu88 of LysF1, is displayed as a stick representation. The similarly positioned residue of T4L, Met102, is shown. (B) The circular dichroism spectra of each mutant compared with wild type LysF1. (C) The thermal stability of wild type LysF1 and each mutant as measured by differential scanning fluorimetry is shown. (D) The muralytic activity of wild type LysF1 as measured in a turbidity reduction assay of outer membrane-permeabilized *E. coli* cells compared with the three Glu88 mutants, Glu88Leu, Glu88Met and Glu88Phe. Activity is reported relative to wild type activity.

Evolutionary conservation of Glu88 amongst homologous proteins was searched for using CONSURF [81]. HMMER sequence similarity search identified 2666 unique hits, and we performed an alignment with 150 sequences. Only one other sequence was observed with a Glu in a similar sequence position. Generally, residues observed in the same positions were hydrophobic such as leucine and phenylalanine (Tables S1 and S2).

We investigated the impact of Glu88 for activity and thermal stability by overexpressing and purifying three variants of LysF1 with the point mutations: Glu88Leu, Glu88Phe and Glu88Met. Leucine and phenylalanine mutations were selected based upon their observed conservation from CONSURF. The methionine mutant was selected for its equivalent position in T4L and was previously found to be more stabilizing than a leucine [78,82], potentially because methionine can occupy the most sidechain conformations and therefore is more malleable within the structure. Initial in silico analysis with STRUM calculated that these point mutations would result in a more stable protein ($\Delta\Delta G$ of -0.68 for Glu88Leu, -1.03 for Glu88Phe and -0.7 for Glu88Met) [83]. Experimentally, we first determined that each mutation did not significantly affect the secondary structure using circular dichroism (Figure 7B). The introduction of a hydrophobic residue in all three variants conferred an increase of thermal stability as measured by differential scanning fluorimetry (Figure 7C). Wild type LysF1 had a melting temperature of 53.6 °C. Glu88Leu was the most stabilizing

mutation, increasing the melting temperature by 9.4 °C. The methionine and phenylalanine mutations increased stability by 7.1 °C and 4.0 °C, respectively.

The impact of each mutation on activity was assessed by a turbidity reduction assay on chloroform-treated *E. coli* cells. The activity of each variant was compared with the wild type (Figure 7D). Variants Glu88Met and Glu88Leu displayed an increase in activity, 123% and 120%, respectively. However, the phenylalanine mutation reduced the activity of LysF1 to 31% of the wild type. We suggest that the reduced flexibility of the phenylalanine sidechain compared with Met, Leu and Glu could hinder the dynamics of the protein, which is likely an important property for binding and hydrolysis of the peptidoglycan.

Minor modification to the amino acid sequence of LysF1 improved the thermal stability and increased activity. LysF1 Glu88Met and Glu88Leu variants are more stable and more active than the wild type endolysin derived from FAHEc1. These findings may be useful for the genetic engineering of FAHEc1. Modifying the genome of FAHEc1 to encode one of these variants could result in a phage that is able to hydrolyze the bacterial peptidoglycan faster which may increase the rate of lysis to be more effective for the control of foodborne pathogens. Alternatively, inclusion of a phenylalanine to modulate activity of the endolysin could be an effective strategy to reduce the inflammatory response if this phage is deployed in a health context [1].

The rational engineering of endolysins, such as site-specific mutagenesis, to enhance their antibacterial activity and stability will be necessary to optimize implementation in various settings [84]. The structural characterization of many more endolysins to elucidate the sequence–structure relationship and identification of key residues will be important to support this. For example, in this study we have identified that a phenylalanine positioned in the hydrophobic core of this protein fold is deactivating and less stable compared with leucine or methionine variants. Phenylalanine is observed in similar sequences and proteins to LysF1, including AcLys [46]. AcLys displayed intrinsic moderate antibacterial activity against Gram-negative bacteria and therefore is a useful foundation for developing a novel antibacterial agent; however, AcLys contains a phenylalanine buried within its core which, based on our elucidated understanding, could be mutated to leucine or methionine to create a more stable and more active enzyme.

4. Conclusions

Bacteriophages and encoded endolysins are promising antibacterial candidates. Here, we report the characterization of the endolysin encoded by FAHEc1, LysF1. We confirm that LysF1 degrades peptidoglycan and causes cell lysis of chloroform-treated Gram-negative bacterial cells. Comparative sequence and structural (3D structure determined by protein x-ray crystallography) analysis reveals that despite varied sequence similarities there is a conservation of the protein architecture, illustrating the robustness of the protein fold. Further structural investigation demonstrated an unusually positioned and non-conserved glutamate within the hydrophobic core. Mutating this residue to hydrophobic amino acids resulted in a more stable, more active protein potentially elucidating an engineering pathway for optimizing similar proteins. Our characterization of LysF1 provides insight into the lytic mechanism of FAHEc1, increasing the structure–function relationship understanding of this class of endolysin, which could be useful in the future for rational engineering and the development of effective endolysin-based antibacterial agents.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/v13061101/s1>, Figure S1: Purification gel of LysF1; Figure S2: Hexagonal prism morphology of LysF1 protein crystals; Figure S3: Residuals of $c(M)$ fit indicate randomly distributed noise; Figure S4: The conserved protein architecture of LysF1. Table S1: Residue conservation for each position of LysF1; Table S2: Sequences used by CONSURF in multisequence alignment.

Author Contributions: M.J.L., D.C., C.B. and R.C.J.D. conceived and designed the experiments. M.J.L., S.H.M. and G.S.A. performed the experiments. M.J.L. and D.C. analyzed the data. M.J.L. wrote the manuscript with input from C.B. and R.C.J.D. All authors have read and agreed to the published version of the manuscript.

Funding: M.J.L. is supported by UC/ESR Connect scholarship. C.B. acknowledges ESR Strategic Science Initiatives Funding. R.C.J.D. acknowledges the following for funding support, in part: (1) the New Zealand Royal Society Marsden Fund (contract UOC1506); (2) a Ministry of Business, Innovation and Employment Smart Ideas grant (contract UOCX1706); and (3) the Biomolecular Interactions Centre (University of Canterbury).

Informed Consent Statement: Not applicable.

Data Availability Statement: The x-ray crystal structure data in this study is available in the Protein Data Bank (PDB) under accession code 7M5L.

Acknowledgments: We would like to thank Michael Currie for support with small angle x-ray scattering data collection, Chris Horne for assistance with sedimentation velocity experiments and Jenna Gilkes for support with x-ray crystallography diffraction data collection. We also thank Brigitta Kurenbach and Jack Heinemann for providing *K. pneumoniae*, *S. typhimurium* and *S. aureus* strains. Parts of this research were undertaken at the MX2 and SAXS/WAXS beamlines of the Australian Synchrotron, Victoria, Australia.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Romero-Calle, D.; Guimarães Benevides, R.; Góes-Neto, A.; Billington, C. Bacteriophages as alternatives to antibiotics in clinical care. *Antibiotics* **2019**, *8*, 138. [[CrossRef](#)] [[PubMed](#)]
2. Chan, B.K.; Abedon, S.T.; Loc-Carrillo, C. Phage cocktails and the future of phage therapy. *Future Microbiol.* **2013**, *8*, 769–783. [[CrossRef](#)] [[PubMed](#)]
3. Loc-Carrillo, C.; Abedon, S.T. Pros and cons of phage therapy. *Bacteriophage* **2011**, *1*, 111–114. [[CrossRef](#)] [[PubMed](#)]
4. Pires, D.P.; Cleto, S.; Sillankorva, S.; Azeredo, J.; Lu, T.K. Genetically engineered phages: A review of advances over the last decade. *Microbiol. Mol. Biol. Rev.* **2016**, *80*, 523–543. [[CrossRef](#)]
5. Furfaro, L.L.; Payne, M.S.; Chang, B.J. Bacteriophage therapy: Clinical trials and regulatory hurdles. *Front. Cell. Infect. Microbiol.* **2018**, *8*, 376. [[CrossRef](#)]
6. Skurnik, M.; Pajunen, M.; Kiljunen, S. Biotechnological challenges of phage therapy. *Biotechnol. Lett.* **2007**, *29*, 995–1003. [[CrossRef](#)] [[PubMed](#)]
7. Love, M.J.; Dobson, R.C.J.; Billington, C. Stemming the tide of antibiotic resistance by exploiting bacteriophages. *Biochemistry* **2020**, *42*, 6–11. [[CrossRef](#)]
8. Roach, D.R.; Donovan, D.M. Antimicrobial bacteriophage-derived proteins and therapeutic applications. *Bacteriophage* **2015**, *5*, e1062590. [[CrossRef](#)]
9. Young, R. Bacteriophage lysis: Mechanism and regulation. *Microbiol. Rev.* **1992**, *56*, 430–481. [[CrossRef](#)]
10. Loessner, M.J. Bacteriophage endolysins—Current state of research and applications. *Curr. Opin. Microbiol.* **2005**, *8*, 480–487. [[CrossRef](#)]
11. Love, M.J.; Bhandari, D.; Dobson, R.C.J.; Billington, C. Potential for bacteriophage endolysins to supplement or replace antibiotics in food production and clinical care. *Antibiotics* **2018**, *7*, 17. [[CrossRef](#)]
12. Fischetti, V.A. Bacteriophage endolysins: A novel anti-infective to control Gram-positive pathogens. *Int. J. Med. Microbiol.* **2010**, *300*, 357–362. [[CrossRef](#)]
13. Nelson, D.C.; Schmelcher, M.; Rodriguez-Rubio, L.; Klumpp, J.; Pritchard, D.G.; Dong, S.; Donovan, D.M. Endolysins as antimicrobials. *Adv. Virus Res.* **2012**, *83*, 299–365. [[CrossRef](#)] [[PubMed](#)]
14. Schmelcher, M.; Donovan, D.M.; Loessner, M.J. Bacteriophage endolysins as novel antimicrobials. *Future Microbiol.* **2012**, *7*, 1147–1171. [[CrossRef](#)]
15. Abdelkader, K.; Gerstmann, H.; Saafan, A.; Dishisha, T.; Briers, Y. The preclinical and clinical progress of bacteriophages and their lytic enzymes: The parts are easier than the whole. *Viruses* **2019**, *11*, 96. [[CrossRef](#)] [[PubMed](#)]
16. Schuch, R.; Khan, B.K.; Raz, A.; Rotolo, J.A.; Wittekind, M. Bacteriophage Lysin CF-301, a potent antistaphylococcal biofilm agent. *Antimicrob. Agents Chemother.* **2017**, *61*. [[CrossRef](#)]
17. Jun, S.Y.; Jang, I.J.; Yoon, S.; Jang, K.; Yu, K.-S.; Cho, J.Y.; Seong, M.-W.; Jung, G.M.; Yoon, S.J.; Kang, S.H. Pharmacokinetics and tolerance of the phage endolysin-based candidate drug SAL200 after a single intravenous administration among healthy volunteers. *Antimicrob. Agents Chemother.* **2017**, *61*, 1–11. [[CrossRef](#)]

18. Totté, J.; de Wit, J.; Pardo, L.; Schuren, F.; van Doorn, M.; Pasmans, S. Targeted anti-staphylococcal therapy with endolysins in atopic dermatitis and the effect on steroid use, disease severity and the microbiome: Study protocol for a randomized controlled trial (MAAS trial). *Trials* **2017**, *18*, 404. [[CrossRef](#)] [[PubMed](#)]
19. Briers, Y.; Lavigne, R. Breaking barriers: Expansion of the use of endolysins as novel antibacterials against Gram-negative bacteria. *Future Microbiol.* **2015**, *10*, 377–390. [[CrossRef](#)] [[PubMed](#)]
20. Briers, Y.; Walmagh, M.; Van Puyenbroeck, V.; Cornelissen, A.; Cenens, W.; Aertsen, A.; Oliveira, H.; Azeredo, J.; Verween, G.; Pirnay, J.-P.; et al. Engineered endolysin-based “Artilynsins” to combat multidrug-resistant Gram-negative pathogens. *MBio* **2014**, *5*, 1–10. [[CrossRef](#)] [[PubMed](#)]
21. Hudson, J.A.; Billington, C.; Cornelius, A.J.; Wilson, T.; On, S.L.W.; Premaratne, A.; King, N.J. Use of a bacteriophage to inactivate *Escherichia coli* O157: H7 on beef. *Food Microbiol.* **2013**, *36*, 14–21. [[CrossRef](#)]
22. Hudson, J.A.; Billington, C.; Wilson, T.; On, S.L.W. Effect of phage and host concentration on the inactivation of *Escherichia coli* O157:H7 on cooked and raw beef. *Food Sci. Technol. Int.* **2015**, *21*, 104–109. [[CrossRef](#)] [[PubMed](#)]
23. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped BLAST and PSI-BLAST: A new generation of protein data base search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)] [[PubMed](#)]
24. Robert, X.; Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **2014**, *42*, W320–W324. [[CrossRef](#)] [[PubMed](#)]
25. Katoh, K.; Misawa, K.; Kuma, K.; Miyata, T. MAFFT: A novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [[CrossRef](#)]
26. Gasteiger, E.; Hoogland, C.; Gattiker, A.; Duvaud, S.; Wilkins, M.R.; Appel, R.D.; Bairoch, A. Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*; Humana Press: Totowa, NJ, USA, 2005; Volume 50, pp. 571–607.
27. Lavigne, R.; Briers, Y.; Hertveldt, K.; Robben, J.; Volckaert, G. Identification and characterization of a highly thermostable bacteriophage lysozyme. *Cell. Mol. Life Sci.* **2004**, *61*, 2753–2759. [[CrossRef](#)]
28. Evans, P.R.; Murshudov, G.N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2013**, *69*, 1204–1214. [[CrossRef](#)]
29. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.W.; McCoy, A.; et al. Overview of the CCP 4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 235–242. [[CrossRef](#)]
30. Stein, N. CHAINSAW: A program for mutating pdb files used as templates in molecular replacement. *J. Appl. Crystallogr.* **2008**, *41*, 641–643. [[CrossRef](#)]
31. McCoy, A.J.; Grosse-Kunstleve, R.W.; Adams, P.D.; Winn, M.D.; Storoni, L.C.; Read, R.J. Phaser crystallographic software. *J. Appl. Crystallogr.* **2007**, *40*, 658–674. [[CrossRef](#)]
32. Murshudov, G.N.; Skubák, P.; Lebedev, A.A.; Pannu, N.S.; Steiner, R.A.; Nicholls, R.A.; Winn, M.D.; Long, F.; Vagin, A.A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 355–367. [[CrossRef](#)] [[PubMed](#)]
33. Emsley, P.; Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60*, 2126–2132. [[CrossRef](#)] [[PubMed](#)]
34. Afonine, P.V.; Grosse-Kunstleve, R.W.; Echols, N.; Headd, J.J.; Moriarty, N.W.; Mustyakimov, M.; Terwilliger, T.C.; Urzhumtsev, A.; Zwart, P.H.; Adams, P.D. Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2012**, *68*, 352–367. [[CrossRef](#)]
35. Laue, T.M.; Shah, B.D.; Ridgeway, T.M.; Pelletier, S.L. Computer-aided interpretation of analytical sedimentation data for proteins. In *Analytical Ultracentrifugation in Biochemistry and Polymer Science*; Harding, S.E., Rowe, A.J., Horton, J.C., Eds.; The Royal Society of Chemistry: Cambridge, UK, 1992; pp. 90–125.
36. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **2000**, *78*, 1606–1619. [[CrossRef](#)]
37. Petoukhov, M.V.; Konarev, P.V.; Kikhney, A.G.; Svergun, D.I. ATSAS 2.1—Towards automated and web-supported small-angle scattering data analysis. *J. Appl. Crystallogr.* **2007**, *40*, 223–228. [[CrossRef](#)]
38. Konarev, P.V.; Volkov, V.V.; Sokolova, A.V.; Koch, M.H.J.; Svergun, D.I. PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* **2003**, *36*, 1277–1282. [[CrossRef](#)]
39. Svergun, D.I. Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* **1992**, *25*, 495–503. [[CrossRef](#)]
40. Šali, A.; Blundell, T.L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779–815. [[CrossRef](#)]
41. Webb, B.; Sali, A. Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinforma.* **2016**, *2016*, 5.6.1–5.6.37. [[CrossRef](#)]
42. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)]
43. Svergun, D.; Barberato, C.; Koch, M.H.J. CRY SOL—A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **1995**, *28*, 768–773. [[CrossRef](#)]

44. Wohlkönig, A.; Huet, J.; Looze, Y.; Wintjens, R. Structural relationships in the lysozyme superfamily: Significant evidence for glycoside hydrolase signature motifs. *PLoS ONE* **2010**, *5*, e15388. [[CrossRef](#)] [[PubMed](#)]
45. Boratyn, G.M.; Schäffer, A.A.; Agarwala, R.; Altschul, S.F.; Lipman, D.J.; Madden, T.L. Domain enhanced lookup time accelerated BLAST. *Biol. Direct* **2012**, *7*, 1–14. [[CrossRef](#)]
46. Sykilinda, N.; Nikolaeva, A.; Shneider, M.; Mishkin, D.; Patutin, A.; Popov, V.; Boyko, K.; Klyachko, N.; Miroshnikov, K. Structure of an *Acinetobacter* broad-range prophage endolysin reveals a C-terminal α -helix with the proposed role in activity against live bacterial cells. *Viruses* **2018**, *10*, 309. [[CrossRef](#)] [[PubMed](#)]
47. Babu, K.; Arulandu, A.; Sankaran, K. The structure of DLP12 endolysin exhibiting alternate loop conformation and comparative analysis with other endolysins. *Proteins Struct. Funct. Bioinforma.* **2018**, *86*, 210–217. [[CrossRef](#)]
48. Mooers, B.H.M.; Matthews, B.W. Extension to 2268 atoms of direct methods in the *ab initio* determination of the unknown structure of bacteriophage P22 lysozyme. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, *62*, 165–176. [[CrossRef](#)] [[PubMed](#)]
49. Sun, Q.; Kutry, G.F.; Arockiasamy, A.; Xu, M.; Young, R.; Sacchettini, J.C. Regulation of a muralytic enzyme by dynamic membrane topology. *Nat. Struct. Mol. Biol.* **2009**, *16*, 1192–1194. [[CrossRef](#)]
50. Xu, M.; Arulandu, A.; Struck, D.K.; Swanson, S.; Sacchettini, J.C.; Young, R. Disulfide isomerization after membrane release of its SAR domain activates P1 lysozyme. *Science* **2005**, *307*, 113–117. [[CrossRef](#)]
51. Weaver, L.H.; Matthews, B.W. Structure of bacteriophage T4 lysozyme refined at 1.7 Å resolution. *J. Mol. Biol.* **1987**, *193*, 189–199. [[CrossRef](#)]
52. Randich, A.M.; Kysela, D.T.; Morlot, C.; Brun, Y.V. Origin of a core bacterial gene via co-option and detoxification of a phage lysin. *Curr. Biol.* **2019**, *29*, 1634–1646.e6. [[CrossRef](#)]
53. Kuroki, R.; Weaver, L.; Matthews, B. A covalent enzyme-substrate intermediate with saccharide distortion in a mutant T4 lysozyme. *Science* **1993**, *262*, 2030–2033. [[CrossRef](#)] [[PubMed](#)]
54. Young, R. Phage lysis: Three steps, three choices, one outcome. *J. Microbiol.* **2014**, *52*, 243–258. [[CrossRef](#)] [[PubMed](#)]
55. Savva, C.G.; Dewey, J.S.; Moussa, S.H.; To, K.H.; Holzenburg, A.; Young, R. Stable micron-scale holes are a general feature of canonical holins. *Mol. Microbiol.* **2014**, *91*, 57–65. [[CrossRef](#)]
56. Walmagh, M.; Boczkowska, B.; Grymonprez, B.; Briers, Y.; Drulis-Kawa, Z.; Lavigne, R. Characterization of five novel endolysins from Gram-negative infecting bacteriophages. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 4369–4375. [[CrossRef](#)]
57. Rodríguez-Rubio, L.; Gerstmanns, H.; Thorpe, S.; Mesnage, S.; Lavigne, R.; Briers, Y. DUF3380 domain from a *Salmonella* phage endolysin shows potent N-acetylmuramidase activity. *Appl. Environ. Microbiol.* **2016**, *82*, 4975–4981. [[CrossRef](#)] [[PubMed](#)]
58. Maciejewska, B.; Żrubek, K.; Espaillet, A.; Wiśniewska, M.; Rembacz, K.P.; Cava, F.; Dubin, G.; Drulis-Kawa, Z. Modular endolysin of *Burkholderia* AP3 phage has the largest lysozyme-like catalytic subunit discovered to date and no catalytic aspartate residue. *Sci. Rep.* **2017**, *7*, 1–12. [[CrossRef](#)]
59. Schleifer, K.H.; Kandler, O. Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriol. Rev.* **1972**, *36*, 407–477. [[CrossRef](#)]
60. McClelland, M.; Florea, L.; Sanderson, K.; Clifton, S.W.; Parkhill, J.; Churcher, C.; Dougan, G.; Wilson, R.K.; Miller, W. Comparison of the *Escherichia coli* K-12 genome with sampled genomes of a *Klebsiella pneumoniae* and three *Salmonella enterica* serovars, Typhimurium, Typhi and Paratyphi. *Nucleic Acids Res.* **2000**, *28*, 4974–4986. [[CrossRef](#)] [[PubMed](#)]
61. Oliveira, H.; Thiagarajan, V.; Walmagh, M.; Sillankorva, S.; Lavigne, R.; Neves-Petersen, M.T.; Kluskens, L.D.; Azeredo, J. A thermostable *Salmonella* phage endolysin, Lys68, with broad bactericidal properties against gram-negative pathogens in presence of weak acids. *PLoS ONE* **2014**, *9*, e108376. [[CrossRef](#)]
62. Krissinel, E.; Henrick, K. Inference of macromolecular assemblies from crystalline State. *J. Mol. Biol.* **2007**, *372*, 774–797. [[CrossRef](#)]
63. Putnam, C.D. Guinier peak analysis for visual and automated inspection of small-angle X-ray scattering data. *J. Appl. Cryst.* **2016**, *49*, 1412–1419. [[CrossRef](#)]
64. Putnam, C.D.; Hammel, M.; Hura, G.L.; Tainer, J.A. X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.* **2007**, *40*, 191–285. [[CrossRef](#)]
65. Combs, C.F.A. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. *Biopolymers* **2011**, *95*, 559–571. [[CrossRef](#)]
66. Anderson, W.F.; Grütter, M.G.; Remington, S.J.; Weaver, L.H.; Matthews, B.W. Crystallographic determination of the mode of binding of oligosaccharides to T4 bacteriophage lysozyme: Implications for the mechanism of catalysis. *J. Mol. Biol.* **1981**, *147*, 523–543. [[CrossRef](#)]
67. Hardy, L.W.; Poteete, A.R. Reexamination of the role of Asp20 in catalysis by bacteriophage T4 lysozyme. *Biochemistry* **1991**, *30*, 9457–9463. [[CrossRef](#)] [[PubMed](#)]
68. Love, M.J.; Abeysekera, G.S.; Muscroft-Taylor, A.C.; Billington, C.; Dobson, R.C.J. On the catalytic mechanism of bacteriophage endolysins: Opportunities for engineering. *Biochim. Biophys. Acta Proteins Proteom.* **2020**, *1868*, 140302. [[CrossRef](#)] [[PubMed](#)]
69. Wang, Q.; Graham, R.W.; Trimbur, D.; Warren, R.A.J.; Withers, S.G. Changing enzymic reaction mechanisms by mutagenesis: Conversion of a retaining glucosidase to an inverting enzyme. *J. Am. Chem. Soc.* **1994**, *116*, 11594–11595. [[CrossRef](#)]
70. McCarter, J.D.; Withers, G.S. Mechanisms of enzymatic glycoside hydrolysis. *Curr. Opin. Struct. Biol.* **1994**, *4*, 885–892. [[CrossRef](#)]
71. Kuroki, R.; Weaver, L.H.; Matthews, B.W. Structural basis of the conversion of T4 lysozyme into a transglycosidase by reengineering the active site. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 8949–8954. [[CrossRef](#)]

72. Wang, S.; Ma, J.; Peng, J.; Xu, J. Protein structure alignment beyond spatial proximity. *Sci. Rep.* **2013**, *3*, 1448. [[CrossRef](#)] [[PubMed](#)]
73. Wang, S.; Peng, J.; Xu, J. Alignment of distantly related protein structures: Algorithm, bound and implications to homology modeling. *Bioinformatics* **2011**, *27*, 2537–2545. [[CrossRef](#)] [[PubMed](#)]
74. Rennell, D.; Bouvier, S.E.; Hardy, L.W.; Poteete, A.R. Systematic mutation of bacteriophage T4 lysozyme. *J. Mol. Biol.* **1991**, *222*, 67–88. [[CrossRef](#)]
75. Starr, T.N.; Thornton, J.W. Epistasis in protein evolution. *Protein Sci.* **2016**, *25*, 1204–1218. [[CrossRef](#)]
76. Isom, D.G.; Castañeda, C.A.; Cannon, B.R.; Velu, P.D.; García-Moreno, E.B. Charges in the hydrophobic interior of proteins. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 16096–16100. [[CrossRef](#)]
77. Kajander, T.; Kahn, P.C.; Passila, S.H.; Cohen, D.C.; Lehtiö, L.; Adolfsen, W.; Warwicker, J.; Schell, U.; Goldman, A. Buried Charged surface in proteins. *Structure* **2000**, *8*, 1203–1214. [[CrossRef](#)]
78. Hurley, J.H.; Baase, W.A.; Matthews, B.W. Design and structural analysis of alternative hydrophobic core packing arrangements in bacteriophage T4 lysozyme. *J. Mol. Biol.* **1992**, *224*, 1143–1159. [[CrossRef](#)]
79. Baase, W.A.; Liu, L.; Tronrud, D.E.; Matthews, B.W. Lessons from the lysozyme of phage T4. *Protein Sci.* **2010**, *19*, 631–641. [[CrossRef](#)]
80. Dao-pin, S.; Anderson, D.E.; Baase, W.A.; Dahlquist, F.W.; Matthews, B.W. Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme. *Biochemistry* **1991**, *30*, 11521–11529. [[CrossRef](#)]
81. Ashkenazy, H.; Abadi, S.; Martz, E.; Chay, O.; Mayrose, I.; Pupko, T.; Ben-Tal, N. ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **2016**, *44*, W344–W350. [[CrossRef](#)] [[PubMed](#)]
82. Lipscomb, L.A.; Gassner, N.C.; Snow, S.D.; Eldridge, A.M.; Baase, W.A.; Drew, D.L.; Matthews, B.W. Context-dependent protein stabilization by methionine-to-leucine substitution shown in T4 lysozyme. *Protein Sci.* **1998**, *7*, 765–773. [[CrossRef](#)]
83. Quan, L.; Lv, Q.; Zhang, Y. STRUM: Structure-based prediction of protein stability changes upon single-point mutation. *Bioinformatics* **2016**, *32*, 2936–2946. [[CrossRef](#)] [[PubMed](#)]
84. São-José, C. Engineering of phage-derived lytic enzymes: Improving their potential as antimicrobials. *Antibiotics* **2018**, *7*, 29. [[CrossRef](#)] [[PubMed](#)]